

What is claimed is:

- 1 1. A system for managing a plurality of computer resource usage quotas, comprising:
- 2 a) a distributed gateway service;
- 3 b) a plurality of clients for accessing the system, said clients being connected to said
- 4 distributed gateway service;
- 5 c) an internal network;
- 6 d) a distributed storage service;
- 7 e) a system management service (SMS);
- 8 f) a configuration database (CDB) for storing the plurality of data storage resource
- 9 usage quotas, said CDB being connected to said SMS; and
- 10 g) a distributed metadata service (MDS) connected to said SMS and said CDB.
- 1 2. The system as defined in Claim 1, wherein the distributed gateway service includes a
- 2 plurality of gateway service nodes.
- 1 3. The system as defined in Claim 1, wherein the plurality of clients comprise one of at least
- 2 one of a NFS client, a CIFS client, a HTTP client, a FTP client, and an SQL client.
- 1 4. The system as defined in Claim 1, wherein the internal network is a non-blocking
- 2 switched network.

1 5. The system as defined in Claim 1, wherein the non-blocking switched network has
2 redundancy.

1 6. The system as defined in Claim 1, wherein said storage service includes:
2 at least one storage server; and
3 at least one storage device connected to said at least one storage server.

1 7. The system as defined in Claim 6, wherein said at least one storage server includes at
2 least one storage class.

1 8. The system as defined in Claim 1, wherein the said MDS includes at least one metadata
2 partition.

1 9. The system as defined in Claim 8, wherein the at least one metadata partition includes an
2 MDS server and an MDS database.

1 10. The system as defined in Claim 9, wherein the MDS server includes a quota usage server
2 and a quota usage updater.

1 11. The system as defined in Claim 9, wherein said MDS database includes a quota usage
2 database.

1 12. The system as defined in Claim 1, further comprising:
2 a distributed quota aggregation and enforcement service connected to the internal
3 network, the CDB, and the MDS.

1 13. The system as defined in claim 12, wherein the distributed quota aggregation and
2 enforcement service includes at least one service node for aggregating quota usage and quota
3 usage limit information.

09852516-051001

1 14. A method for managing the cost of utilization of computer resources, said method
2 comprising the steps of:

- 3 a) defining classes of computer resources based on quality of service attributes;
4 b) defining computer resource usage policies based on said class definitions;

5 whereby the cost utilization of the computer resources is managed by using the computer
6 resources according to said usage policies.

15. The method as defined in Claim 14 wherein said quality of service attributes are selected
from a group including resource performance, resource reliability, resource availability,
resource latency, and resource bandwidth.

16. The method as defined in Claim 14 wherein said step a) further includes assigning each
class an identification designator.

17. The method as defined in Claim 14 wherein said step b) further includes the step of
assigning usage quotas to each of said defined classes.

18. The system as defined in Claim 17 wherein said usage quotas are hard quotas.

19. The method as defined in Claim 17 wherein said usage quotas are soft quotas.

- 1 20. A method for defining, tracking, and enforcing computer resource usage quota
2 comprising the steps of:
3 a) defining classes;
4 b) defining computer resource usage policies;
5 c) applying said usage policies to said classes;
6 d) generating reports on the usage of said classes;
7 e) using the computer resource according to said defined class in accordance with
8 said computer resource usage policies.
21. The method as defined in Claim 20 wherein said class definitions assign a cost to a defined class.
22. The method as defined in Claim 21 wherein said step a) further includes:
i) defining quality of service attributes.
23. The method as defined in Claim 22 wherein said service attributes are selected from a group including resource performance resource reliability, resource availability, resource latency, and resource bandwidth.
24. The system as defined in Claim 23 wherein each class is assigned a class identification.

1 25. The method as defined in Claim 20 wherein said step b) further includes:

2 i) assigning usage quotas to said defined classes.

1 26. The method as defined in Claim 20 wherein said step c) further includes:

2 i) tracking, aggregating, and enforcing said computer resource quotas;

1 27. The method as defined in Claim 20 wherein said quotas defined in step b) are hard
2 quotas.

28. The method as defined in Claim 27 wherein said hard quota cannot be exceeded under
any circumstances.

29. The method as defined in Claim 20 wherein quotas defined in step b) are soft quotas.

30. The method as defined in Claim 29 wherein said soft quota may be exceeded for a
predetermined period of time.

1 31. The method as defined in Claim 26 wherein said enforcement includes sending
2 notification when said resource usage quota has been exceeded.

1 32. The method as defined in Claim 31 further including rejecting or accepting a request to
2 use a computer resource when a quota has been exceeded.